MAR - 2 1977

NEW YORK UNIVERSITY
WASHINGTON SQUARE COLLEGE OF ARTS AND SCIENCE
MATHEMATICS RESEARCH GROUP

RESEARCH REPORT No. TW-16

EIGENVALUES OF COMPOUND MATRICES

by

BERNARD FRIEDMAN

The research reported in this document has been made possible through support and sponsorship extended by the Air Force Cambridge Research Center, AMC, under Contract No. AF-19(122)-1. It is published for technical information only and does not represent recommendations or conclusions of the sponsoring agency.

October, 1951

c.2

NEW YORK UNIVERSITY
WASHINGTON SQUARE COLLEGE OF ARTS AND SCIENCE
MATHEMATICS RESEARCH GROUP
RESEARCH REPORT No. TW-16

EIGENVALUES OF COMPOUND MATRICES

by

BERNARD FRIEDMAN

Written by:

Bernard Friedman
Bernard Friedman

Morris Kline
Morris Kline
Project Director

Title page
16 numbered pages
October, 1951

TABLE OF CONTENTS

Section	Page
1. Introduction	1
2. The Direct Product of Matrices	3
3. Permutation Matrices	7
4. Applications	11
5. Condition of a Matrix	13
Bibliography	16

ABSTRACT

A method is proposed to simplify the eigenvalue problem for matrices whose elements are matrices. The method is useful in cases where the original matrix has a group of symmetries. Applications of the method to wave guide functions, the excluded volume effect in polymers, the Ising model of ferromagnetism and to numerical analysis are given.

1. Introduction

The theory of wave guide junctions and of directional couplers in particular employs matrices to represent scattering operators and symmetry operators. In the application of this theory it is necessary to find the eigenvalues of fourth, sixth, and, potentially, even higher order matrices. Because these matrices generally possess some special symmetry properties it should be possible to simplify the computation of such eigenvalues. An investigation was made of the problem of calculating the eigenvalues of certain compound matrices which are described more fully below. The theory of this paper includes those cases encountered in wave guide practice and makes feasible more complex applications than have hitherto been attempted, as, for example, a junction of junctions. Several other applications of the theory are also presented.

This paper discusses the eigenvalues of nr-rowed square matrices which can be written as r-rowed square matrices with elements which are n-rowed square matrices. Of course, every nr-rowed square matrices can be written in this way and in general nothing can be said about the eigenvalues. If, however, the r-rowed matrix possesses some kind of symmetry (in a sense to be made precise later) we obtain results which show that the eigenvalues of the original matrix can be expressed in terms of eigenvalues of the n-rowed matrix elements. The simplest example of such a result is the following:

Let A and B be n-rowed square matrices and consider the following 2n-rowed matrix:

$$\begin{pmatrix} A & B \\ B & A \end{pmatrix}$$

The eigenvalues of this matrix are the same as the eigenvalues of the two n-rowed matrices, A + B and A - B. This means that instead of solving a $2n^{th}$-degree determinantal equation we need solve two n^{th}-degree equations.

The results we obtain are based on the notion of the direct product of two matrices [1] . In Section 2, this notion is explained and we use it to prove Theorems I and II which show how the sum of direct products of matrices can be simplified. These results generalize theorems of D. E. Rutherford [2] and of Burnside [3].

In Section 3 we consider matrices such that each row is obtained from the preceding by a cyclic permutation in which the elements (n-rowed matrices) are moved p-places to the right. We show how the eigenvalue problem for such a matrix can be reduced to the consideration of the eigenvalue problem for sub-matrices. The number and the degree of the matrices will depend on the coset decomposition of the multiplicative semi-group of integers (mod r) with respect to the cyclic group generated by p.

In Section 4 three applications of the theorems to matrices which occur in physical problems are discussed. These applications are taken from the theory of directional couplers of wave guides [4], [8], the theory of overlapping polymer chains [5] and the Ising model of ferromagnetic materials [6].

Section 5 is an application to numerical analysis. Van Neumann and Goldstone [7] have defined the "condition" of a matrix, which is a measure of how well it is adapted for numerical computation. We show that the condition of the matrix for a finite difference approximation to the Laplacean in n-dimensions equals the condition of the same finite difference approximation to the ordinary second derivative in one dimension.

2. The Direct Product of Matrices.

Consider two linear spaces S_n and S_r of dimensions n and r respectively. Let ξ be an arbitrary vector in S_n with components $\xi_1, \xi_2, \ldots \xi_n$ and let η be an arbitrary vector in S_r with components $\eta_1, \eta_2 \ldots \eta_r$. The direct product of ξ and η is defined to be the vector ζ with nr components $\xi_i \, \eta_j (i = 1,2,\ldots n; \; j = 1,2,\ldots r)$. We write $\zeta = \xi \times \eta$. The linear space spanned by all possible vectors ζ is called the direct product of the spaces S_n and S_r. It is written $S_n \times S_r$ and is obviously of dimension nr.

Suppose A is an n-dimensional matrix with elements $a_{i'i}$ acting on the vectors ξ and suppose B is an r-dimensional matrix with elements $b_{j'j}$ acting on the vectors η ; then the direct product [1] of A and B (written AxB) is the transformation C in nr dimensions which transforms the vectors ζ as follows:

$$(1) \quad C\zeta^{(1)} = A\xi^{(1)} B\eta^{(1)} = \sum_i \sum_j a_{i'i} b_{j'j} \, \xi_i^{(1)} \eta_j^{(1)} = \xi_{i'}^{(2)} \times \eta_{j'}^{(2)} = \zeta^{(2)}$$

If the nr components of ζ are ordered in some way, then C will be represented by a matrix. To different orderings of the components will correspond different matrices. For example, put

$$z_k = \xi_i \, \eta_j$$

where

a) $k = (i-1)r + j$

or

b) $k = i + (j-1)n$.

Note that any other relation which assigns a unique number k, $(1 \leq k \leq nr)$ to the pair of numbers i, j, $(1 \leq i \leq n, \; 1 \leq j \leq r)$ can be used. The matrix C has elements $c_{k'k}$ where

$$(2) \qquad c_{k'k} = a_{i'i} b_{j'j} \, .$$

Here k' bears the same relation to the pair (i',j') that k bears to the pair (i,j).

If relation a) is used, then C is represented by the following matrix:

$$(3) \qquad \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{n1}B & a_{n2}B & & a_{nn}B \end{pmatrix}$$

However, if relation b) is used, then C is represented by the following:

$$(4)\qquad \begin{pmatrix} b_{11}A & b_{12}A & \cdots & b_{1n}A \\ b_{21}A & b_{22}A & \cdots & b_{2n}A \\ \vdots & \vdots & & \vdots \\ b_{n1}A & b_{n2}A & & b_{nn}A \end{pmatrix}$$

Even though (3) and (4) are not the same matrices, they represent the same transformation but referred to different sets of coordinate axes.

From the definition of direct product and from (2) it is easy to see that the direct product of matrices obeys the following rules [1]:

1) $$(A_1 + A_2) \times B = A_1 \times B + A_2 \times B$$

where the addition is the ordinary matrix addition;

2) $$(A_1 \times B_1) \cdot (A_2 \times B_2) = (A_1 \cdot A_2) \times (B_1 \cdot B_2)$$

where the multiplication is the ordinary matrix multiplication.

Suppose that λ is an eigenvalue of A and μ an eigenvalue of B, then it is well known [1] that $\lambda\mu$ is an eigenvalue of A x B; for, let ξ_i be the eigenvector of A corresponding to the eigenvalue λ and let η_j be the eigenvector of B corresponding to the eigenvalue μ, then $\zeta = \xi_i \eta_j$ is an eigenvector of A x B since

$$(A \times B)\zeta = \sum \sum a_{i'i} b_{j'j} \xi_i \eta_j = \sum a_{i'i} \xi_i \sum b_{j'j} \eta_j = \lambda \xi_{i'} \mu \eta_{j'}$$

$$= \lambda\mu \xi_{i'} \eta_{j'} .$$

A useful generalization of this result is given by

Theorem I. Let $A_1, A_2, \ldots, A_p$ be n-dimensional square matrices and $B_1, B_2 \ldots, B_p$ be r-dimensional square matrices. Suppose that $B_1, B_2, \ldots B_p$ have a common eigenvector, η, and that the corresponding eigenvalues are $\mu_1, \mu_2, \ldots \mu_p$ respectively; then the eigenvalues of

$$(5)\qquad \mu_1 A_1 + \mu_2 A_2 + \ldots + \mu_p A_p$$

will be eigenvalues of

$$(6)\qquad C = A_1 \times B_1 + A_2 \times B_2 + \ldots + A_p \times B_p$$

Proof:

Let ξ be an eigenvector of (5) corresponding to the eigenvalue λ so that

$$(\mu_1 A_1 + \mu_2 A_2 + \ldots + \mu_p A_p)\,\xi = \lambda \xi . \tag{7}$$

From the definition of η, we have

$$B_1 \eta = \mu_1 \eta,\ B_2 \eta = \mu_2 \eta,\ \ldots\ ,\ B_p \eta = \mu_p \eta;$$

then by (7) it follows that

$$\begin{aligned}(A_1 \times B_1 + \ldots + A_p \times B_p)\,\xi \times \eta &= (A_1 \xi) \times (B_1 \eta) + \ldots + (A_p \xi) \times (B_p \eta)\\ &= (A_1 \xi) \times (\mu_1 \eta) + \ldots + (A_p \xi) \times (\mu_p \eta)\\ &= (\mu_1 A_1 \xi + \ldots \mu_p A_p \xi) \times \eta = \lambda \xi \times \eta .\end{aligned}$$

This shows that $\xi \times \eta$ is an eigenvector of (6) corresponding to the eigenvalue λ.

It is well known [1] that a set of commuting matrices have a common eigenvector. This suggests the following

Corollary 1. If the matrices B_1, B_2, ... B_p all commute, then the conclusions of Theorem 1 hold.

A simple consequence of this corollary is

Corollary 2. Let I_n be the unit matrix in n dimensional space and I_r the unit matrix in r dimensional space, then the eigenvalues of

$$A \times I_r + I_n \times B$$

are, $\lambda_i + \mu_j$, where λ_i is any eigenvalue of A and μ_j is any eigenvalue of B.

Corollary 2 is identical with a theorem given by D.E. Rutherford on page 23 of reference |2|. We shall illustrate Theorem I by considering the following matrix:

$$C = \begin{pmatrix} A_1 & A_2 & A_3 \\ A_3 & A_1 & A_2 \\ A_2 & A_3 & A_1 \end{pmatrix} \tag{8}$$

where A_1, A_2, A_3 are n-dimensional square matrices. We may write

$$C = A_1 \times B_1 + A_2 \times B_2 + A_3 \times B_3$$

where

$$B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\quad B_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},\quad B_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Since $B_2^3 = B_1$, the identity matrix, the eigenvalues of B_2 are 1, ρ, ρ^2 where ρ is a cube root of unity. Let η_0, η_1, η_2 be the eigenvectors of B_2 corresponding to the eigenvalues 1, ρ, ρ^2 respectively, then since $B_3 = B_2^2$, we have

$$B_3 \eta_0 = \eta_0, \quad B_3 \eta_1 = \rho^2 \eta_1, \quad B_3 \eta_2 = \rho \eta_2$$

so that η_0, η_1, η_2 are eigenvectors for all three matrices B_1, B_2, and B_3. From Theorem I we conclude that the eigenvalues of C a matrix of 3n dimensions will be given by eigenvalues of the three n-dimensional matrices:

$$A_1 + A_2 + A_3, \quad A_1 + \rho A_2 + \rho^2 A_3, \quad A_1 + \rho^2 A_2 + \rho A_3$$

Theorem I may be generalized to cases where B_1, B_2, ..., B_p may not have a common eigenvector. Consider the ring, R, generated by the matrices B_1, B_2,...,B_p under matrix addition and multiplication. The ring, R, considered as an abstract ring, has an r-dimensional representation given by the matrices B_1, B_2, ..., B_p and the matrices generated from them. If this r-dimensional representation is irreducible, nothing can be said about C. On the other hand, if the r-dimensional representation is reducible, it can be written as the direct sum of irreducible representations. Every such representation of degree one implies that the matrices B_1, B_2, ..., B_p have a common eigenvector. A representation of degree m implies that there exists an m-dimensional subspace which is invariant under the transformations of the ring R. Let $\eta_j^{(1)}$, $\eta_j^{(2)}$, ..., $\eta_j^{(m)}$, $j = 1, \ldots, r$, be the components of a basis for this subspace;,then, if $b_{j'j}^{(k)}$, $k = 1, \ldots, p; j, j' = 1, \ldots, r$, are the elements of a matrix B_k, we have

$$(9) \qquad \sum_j b_{j'j}^{(k)} \eta_j^{(h)} = \sum_{h'} m_{h'h}^{(k)} \eta_{j'}^{(h')}, \qquad h = 1, 2, \ldots, m$$

so that $m_{h'h}^{(k)}$ are the elements of a matrix, M_k, which represents B_k.

We now state

Theorem II. Let C be the matrix considered in Theorem I. Suppose that the ring generated by the matrices B_1, ... B_p has an m-dimensional representation ($m<r$) in which the matrix B_k is represented by M_k; then every eigenvalue of the nm dimensional matrix

$$D = A_1 \times M_1 + \ldots + A_p \times M_p$$

is an eigenvalue of the nr-dimensional matrix

$$C = A_1 \times B_1 + \ldots + A_p \times B_p \quad .$$

Proof:

Let $a^{(k)}_{i'i}$, $i = 1, 2, \ldots, n$, be the elements of the matrix A_k and suppose ξ_{ih}, $i = 1, 2, \ldots, n$; $h = 1, 2, \ldots, m$, is an eigenvector of the matrix D so that

$$\sum_k \sum_h \sum_i a^{(k)}_{i'i} \, m^{(k)}_{h'h} \, \xi_{ih} = \lambda \, \xi_{i'h'} \quad . \tag{10}$$

We show now that $\sum_h \xi_{ih} \eta^{(h)}_j$ is an eigenvector of C; for, using (9) and (10), we have

$$C \sum_h \xi_{ih} \, \eta^{(h)}_j = \sum_k \sum_j \sum_i a^{(k)}_{i'i} \, b^{(k)}_{j'j} \sum_h \xi_{ih} \, \eta^{(h)}_j \quad ,$$

$$= \sum_k \sum_j \sum_i a^{(k)}_{i'i} \, \xi_{ih} \sum_h b^{(k)}_{j'j} \, \eta^{(h)}_j \quad ,$$

$$= \sum_k \sum_j \sum_i a^{(k)}_{i'i} \, \xi_{ih} \sum_{h'} m^{(k)}_{h'h} \, \eta^{(h')}_{j'} \quad ,$$

$$= \sum_{h'} \eta^{(h')}_{j'} \sum_k \sum_j \sum_i a^{(k)}_{i'i} \, m^{(k)}_{h'h} \, \xi_{ih} \quad ,$$

$$= \lambda \sum_{h'} \xi_{i'h'} \, \eta^{(h')}_{j'} \quad .$$

This also shows that λ is an eigenvalue of C and so completes the proof.

3. Permutation Matrices.

Theorem II can be used to analyze permutation matrices which are of the same type as (8). A matrix will be called a permutation matrix if the elements of any row are a permutation of the elements of the first row. Let $A_1, A_2, \ldots, A_r$ be the elements of the first row of a permutation matrix ξ . Let the elements of the second row be the elements of the first row shifted cyclically p places to the right so that the elements of the second row are $A_{r-p+1}, A_{r-p+2}, \ldots A_r, A_1, A_2, \ldots A_{r-p}$ in that order. Let the elements of the third row be the elements of the first row shifted 2p places to the right and in general let the elements of the kth row be the elements of the first row shifted (k-1) places to the right. Such a matrix will be called a p-cycle permutation matrix. Note that (8) is a one-cycle permutation matrix.

Suppose C is a p-cycle permutation matrix; then it is easy to see that

$$C = A_1 \times B_1 + A_2 \times B_2 + \ldots + A_r \times B_r \tag{11}$$

where $B_1, B_2, \ldots, B_r$ are r-dimensional matrices, all of whose elements are zeros except for one element in each row which is unity. If p is prime to r, there will be just one unity in every row and column so that the matrix B_k will be non-singular In this case the matrices $B_1, \ldots, B_r$ generate a group, G, and by Theorem II the irreducible representations of G can be used to simplify the matrix C.

The irreducible representations of G will depend on the r-th roots of unity and the number-theoretic properties of p, modulo r. Hereafter, we assume that p is prime to r. Let t be the exponent to which p belongs modulo r, that is, t is the smallest positive integer such that

$$(12) \qquad p^t \equiv 1 \pmod{r};$$

then the multiplicative semi-group of integers 1,2,...r-1, modulo r, can be written as a sum of cosets relative to the cyclic group $1, p, p^2, \ldots, p^{t-1}$. In other words, the set of integers 1, 2, ..., r-1, modulo r can be written as $1, p, p^2, \ldots, p^{t-1}$; $a_1, a_1p, a_1p^2, \ldots, a_1p^{t-1}$; $a_2, a_2p, \ldots, a_2p^{t-1}$; and so on, where a_1 is any integer, not divisible by r, which is not in the set $1, p, \ldots, p^{t-1}$ a_2 is any integer not divisible by r, which is not in either of the sets $1, p, \ldots, p^t$ or $a_1, a_1p, \ldots, a_1p^{t-1}$, and so on. Let ρ be a primitive r-th root of unity and define

$$(13) \qquad S_k = A_1 + \rho^k A_2 + \rho^{2k} A_3 + \ldots + \rho^{k(r-1)} A_r, \qquad k \equiv 0, 1, 2, \ldots, r-1, (\text{mod } r);$$

then we shall prove the following:

Theorem III. Let C be a p-cycle permutation matrix; then the eigenvalues of any of the following matrices;

$$(14) \qquad T_0 = (S_0)$$

$$T_1 = \begin{pmatrix} 0 & 0 & \cdot & \cdot & S_{p^{t-1}} \\ S_1 & 0 & \cdot & \cdot & 0 \\ \cdot & S_p & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 0 & \cdot \\ 0 & \cdot & \cdot & S_{p^{t-2}} & 0 \end{pmatrix}, \quad T_{a_1} = \begin{pmatrix} 0 & 0 & \cdot & \cdot & \cdot & S_{a_1p^{t-1}} \\ S_{a_1} & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & S_{a_1p} & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & \cdot \\ 0 & \cdot & \cdot & \cdot & S_{a_1p^{t-2}} & 0 \end{pmatrix}, \ldots$$

will be eigenvalues of C.

As was remarked before, (8) is an example of a one-cycle permutation matrix. Obviously $t = 1$ and then by Theorem III, the eigenvalues of S_0, or S_1, or S_2 will be eigenvalues of (8). This agrees with our previous result.

For another illustration, consider a minus-one cycle permutation matrix with $r = 6$. By a minus-one-cycle matrix is meant a matrix where the elements of each row are the elements of the preceding row moved cyclically one place to the left. Let ρ be a primitive sixth root of unity and let S_k be defined by (13). Note that $S_{-k} = S_{6-k}$ and that $t = 2$; then from Theorem III the eigenvalues of

$$(S_0),\ (S_3),\ \begin{pmatrix} 0 & S_{-1} \\ S_1 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & S_{-2} \\ S_2 & 0 \end{pmatrix},$$

will be eigenvalues of the original matrix.

Proof:

In the proof of Theorem III we shall not try to find the irreducible representations of G. Instead, we shall find a matrix U which transforms the original representation of G into a direct sum of irreducible representations.

Let the elements of C be written as A_{ij} where

$$(15) \qquad A_{ij} = A_{i(r-p)+j+1}, \qquad i,j \equiv 0,1,2,\dots,r-1,(\text{mod } r).$$

Suppose that the integers 1, 2, ..., r-1 are written as the sum of the cosets;

$$(16) \qquad 1,p,p^2,\ \dots,\ p^{t-1};\ a_1,a_1p,\dots,a_1p^{t-1};\ a_2,a_2p,\dots a_2p^{t-1};\ \text{and so on.}$$

Let α_k,$[k \equiv 1, 2, \dots, r-1, (\text{mod } r)]$, be the kth integer in the sequence (16) and let $\alpha_0 \equiv 0$ (mod r).

Define a matrix U with elements u_{jk} as follows; u_{jk} is I_n (the identity matrix in n dimensions) times ρ to the exponent $j\alpha_k$, where $j,k \equiv 0,1,\dots,r-1,(\text{mod } r)$. Let V be a matrix with elements v_{jk} where

$$(17) \qquad v_{jk} = u_{kj}^{-1}\ .$$

Since ρ is a primitive rth root of unity,

$$(18) \qquad \sum_{j=0}^{r-1} \rho^{\alpha j} = 0$$

unless $\alpha \equiv 0$ (mod r) in which case the sum is r. Now, by (17) and (18),

$$VU = (\sum_j v_{ij}u_{jk}) = (\sum u_{ji}^{-1}u_{jk})$$

$$= (I_n \sum_j \rho^{(\alpha_k-\alpha_i)j}) = (r\,\delta_{ik}I_n),\ i,k \equiv 0,1,\ldots,r-1, (\text{mod } r)\ ,$$

which is r times the identity matrix in nr-dimensional space. We see then that

$$r^{-1/2}V = (r^{-1/2}U)^{-1}\ .$$

The eigenvalues of C will be the same as those of the similar matrix

$$C_1 = (r^{-1/2}U)^{-1}C(r^{-1/2}U) = r^{-1}VCU.$$

The elements of C_1 are

$$(19) \qquad r^{-1}\sum_j\sum_k v_{ij}a_{jk}u_{kh} = \sum_j\sum_k \rho^{-j\alpha_i}A_\beta\rho^{k\alpha_h}$$

where by (15)

$$(20) \qquad \beta = j(r-p) + k + 1.$$

For fixed j, β runs over a complete residue system mod r as k varies; hence, summing over β instead of k, we find that the sum in (19) becomes

$$\sum_j \rho^{-j\alpha_i}\sum_\beta A_\beta\rho^{(\beta-1)\alpha_h}\,\rho^{-j(r-p)\alpha_h}$$

$$= \sum_j \rho^{j(p\alpha_h-\alpha_i)}S_{\alpha_h}$$

from (13). Again, by (18) this sum is zero unless

$$\alpha_i \equiv p\alpha_h (\text{mod } r)\ .$$

We see then that the elements of C_1 are

$$(21) \qquad \begin{pmatrix} T_o & & & & & \\ & T_1 & & & & \\ & & T_{a_1} & & & \\ & & & T_{a_2} & & \\ & & & & . & \\ . & & & & & . \end{pmatrix}$$

and from this representation the conclusion of Theorem III follows.

If $A_1, A_2, \ldots, A_r$ are not matrices but ordinary scalars, then the eigenvalues of C can be found immediately. We state the following

Corollary. Let C be a p-cycle permutation matrix and let the elements $A_1, \ldots, A_r$ be scalars; then the S_k are scalars and the eigenvalues of C are S_o, the t-th roots of

$$S_1 \, S_p \, S_{p^2} \ldots S_{p^{t-1}} ,$$

the t-th roots of

$$S_{a_1} \, S_{a_1 p} \ldots S_{a_1 p^{t-1}} ,$$

and so on.

This corollary includes some well known results in determinants [3].

4. Applications.

To indicate the application of the above theory to wave guide junctions we consider the matrix

$$S = \begin{pmatrix} \alpha & \beta & \gamma & \delta \\ \beta & \alpha & \delta & \gamma \\ \gamma & \delta & \alpha & \beta \\ \delta & \gamma & \beta & \alpha \end{pmatrix} = \begin{pmatrix} A & B \\ B & A \end{pmatrix}$$

which appears in the theory of directional couplers (see page 437 reference [4]). To obtain the eigenvalues of this matrix we apply Theorem III and conclude first that the eigenvalues of S are the eigenvalues of

$$A + B = \begin{pmatrix} \alpha + \gamma & \beta + \delta \\ \beta + \delta & \alpha + \gamma \end{pmatrix} \quad \text{and} \quad A - B = \begin{pmatrix} \alpha - \gamma & \beta - \delta \\ \beta - \delta & \alpha - \gamma \end{pmatrix} .$$

We may now apply the theorem again to assert that the eigenvalues of S are the eigenvalues of the one-element matrices

$T_o = (\alpha + \gamma + \beta + \delta)$, $T_1 = (\alpha + \gamma - \beta - \delta)$, $T_3 = (\alpha - \gamma + \beta - \delta)$, $T_4=(\alpha-\gamma-\beta+\delta)$.

Of course the eigenvalue of each T_i is the element itself. This result should be compared with formula (129) on p. 441 of reference [4].

In a recent paper on overlapping polymer chains [5] the eigenvalues of the following matrix had to be found:

$$(23) \qquad \begin{pmatrix} A_1 & A_2 & A_3 & A_4 \\ A_2 & A_1 & A_4 & A_3 \\ A_3 & A_4 & A_1 & A_2 \\ A_4 & A_3 & A_2 & A_1 \end{pmatrix}$$

This matrix it will be noted has the same structure as that just considered but here the elements A_i are matrices instead of scalars. From the nature of the problem it was evident that there existed a symmetry group, namely: the group of reflections about the x or y-axis. Since this group is commutative, all its irreducible representations are one-dimensional and thus by Theorem I or II, the eigenvalues of (23) are the eigenvalues of $A_1+A_2+A_3+A_4$, or $A_1-A_2+A_3-A_4$, or $A_1+A_2-A_3-A_4$, or $A_1-A_2-A_3+A_4$.

In a paper on the Ising model for ferromagnetic materials [5] the eigenvalues of the product of the following two matrices had to be found:

$$(24)\quad \begin{pmatrix} A & & & & & \\ & A & & & & \\ & & A & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & A \end{pmatrix}, \qquad \begin{pmatrix} \cos\theta' & & & & & -\sin\theta' \\ & B & & & & \\ & & B & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & B & \\ \sin\theta' & & & & & \cos\theta' \end{pmatrix}$$

Here A and B are two-dimensional rotation matrices, as follows:

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \qquad B = \begin{pmatrix} \cos\theta' & \sin\theta' \\ -\sin\theta' & \cos\theta' \end{pmatrix}$$

The first matrix of (24) can obviously be written as $A \times I_n$ where I_n is the identi matrix in n-dimensions. The second matrix of (24) is not a direct product but it can be seen to be a sum of direct products if it is written as follows:

$$\begin{pmatrix} \cos\theta' & 0 & 0 & 0 & & & & 0 & -\sin\theta' \\ 0 & \cos\theta' & \sin\theta' & 0 & & & & 0 & 0 \\ 0 & -\sin\theta' & \cos\theta' & 0 & \cdot\ \cdot & & & & \\ 0 & 0 & 0 & \cos\theta' & \cdot\ \cdot & & & & \\ & & \cdot & \cdot & \cdot\ \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot\ \cdot & \cdot & \cdot & & \\ & & & & \cdot\ \cdot & \cos\theta' & 0 & 0 & 0 \\ & & & & \cdot\ \cdot & 0 & \cos\theta' & \sin\theta' & 0 \\ 0 & 0 & & & & 0 & -\sin\theta' & \cos\theta' & 0 \\ \sin\theta' & 0 & & & & 0 & 0 & 0 & \cos\theta' \end{pmatrix}$$

which can be written as

$$(25)\quad \begin{pmatrix} A_2 & A_3 & \cdot & \cdot & \cdot & A_4 \\ A_4 & A_2 & A_3 & \cdot & \cdot & \cdot \\ \cdot & A_4 & A_2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & A_3 \\ A_3 & \cdot & \cdot & \cdot & A_4 & A_2 \end{pmatrix} \qquad = A_2 \times I_n + A_3 \times J_n + A_4 \times K_n$$

where

$$A_2 = \begin{pmatrix} \cos\theta' & 0 \\ 0 & \cos\theta' \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 \\ \sin\theta' & 0 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & -\sin\theta' \\ 0' & 0 \end{pmatrix} \tag{26}$$

$$J_n = \begin{pmatrix} 0 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & \cdot & 0 \end{pmatrix} \qquad K_n = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & 0 \end{pmatrix} \tag{27}$$

The product of the matrices is

$$(A x I_n)\,(A_2 x I_n + A_3 x J_n + A_4 x K_n) \;=\; AA_2 x I_n + AA_3 x J_n + AA_4 x K_n$$

by the rules for operating with direct products. It is easy to show that

$$J_n^{n-1} = K_n \;, \quad J_n^n = I_n$$

so that I_n, J_n and K_n all have a common eigenvector. If ρ is a primitive nth root of unity, then the eigenvalues of I_n, J_n, K_n are 1, ρ^k, ρ^{-k} $(k = 0, 1, \ldots, n-1)$ respectively. From Theorem I we conclude that the eigenvalues of the n two-dimensional matrices

$$AA_2 + \rho^k AA_3 + \rho^{-k} AA_4 \;, \quad k = 0, 1, \ldots, n-1 \tag{28}$$

will be eigenvalues of the original product. This agrees with the results.

5. Condition of a Matrix.

The usefulness for computational purposes of a finite difference approximation to a partial differential equation depends on the matrix of the finite difference equations. One test of whether a matrix is "well-conditioned" for computational purposes depends on the size of its P-condition number, which has been defined as the ratio of the greatest to the least (in absolute value) characteristic root of the matrix [7].

We consider the following partial differential equation:

$$(29) \qquad \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} = \lambda u$$

within a cube, of side h, on the boundary of which u vanishes. Divide the cube, which has vertices at (0, 0, 0) and (h,h, h), into $(n+1)^3$ cubes each of side h by the lines $x_1 = rh_1$, $x_2 = sh$, $x_3 = th$, where r, s, t, = 1, 2, ..., n. Suppose that the second derivative in any direction is replaced by some finite difference approximation involving only points on a line parallel to the direction of differentation. For example, if we put

$$u_{r,s,t} = u(rh,sh,th),$$

a possible approximation to $\frac{\partial^2 u}{\partial x_1^2}$ would be

$$\left\{ u_{r+1,\, s,\, t} - 2u_{r,\, s,\, t} + u_{r-1,\, s,\, t} \right\} h^{-2}.$$

In general the finite difference approximation can be represented by the action of some matrix M on the coordinates along the line of differentation. If we write I for the n x n unit matrix, we see that the finite difference approximation to (29) is

$$(30) \qquad (M \times I \times I + I \times M \times I + I \times I \times M)\, u_{r,\, s,\, t} = \lambda h^2 u_{r,s,t}$$

Now, by Theorem I, the eigenvalues of the matrix on the left side of (30) are

$$\lambda_i + \lambda_j + \lambda_k, \quad i, j, k, = 1, 2, \ldots, n$$

where λ_i, λ_j, λ_k, are eigenvalues of M. It is clear that the P-condition number of the matrix in (30) is

$$3\lambda_{max} / 3\lambda_{min} = \lambda_{max} / \lambda_{min}$$

where λ_{max} is the maximum, λ_{min} the minimum (in absolute value) of the eigenvalues of M; therefore, the P-condition number of the matrix in (30) is the same as the P-condition number of M.

This result can be obviously generalized to more dimensions. We have then the following:

Theorem IV. Let M be the matrix associated with the finite difference approximation to the second derivative. Suppose the same approximation is used in each direction for the partial differential operator:

$$(31)\qquad \frac{\partial^2 u}{\partial x_1^{\,2}} + \frac{\partial^2 u}{\partial x_2^{\,2}} + \dots + \frac{\partial^2 u}{\partial x_n^{\,2}} = \lambda u$$

over an n-dimensional cube; then the P-condition number of the matrix associated with the approximation to (31) is the same as the P-condition number of M.

Bibliography

[1] MacDuffee, C.C., The Theory of Matrices, Chelsea Publishing Co., New York, N.Y., 1946

[2] Rutherford, D.E., Some Continuant Determinants Arising in Physics and Chemistry, Proc. Roy. Soc. Edinburgh, A, vol 62 (1946) p. 229

[3] Burnside, W., Mess Math. II vol.23, (1894) p.112

[4] Montgomery, C.G., Dicke, R.H., and Purcell, E.M., Principles of Microwave Circuits, McGraw-Hill Book Co., Inc., N.Y., 1948, [See Chapter 12 in particular]

[5] Montroll, E.W., Markoff Chains and Excluded Volume Effect in Polymer Chains, Journ. of Chem. Phys., vol. 18, 1950, p.734

[6] Kaufman, B., Crystal Statistics II, Phys. Rev. vol. 76 (1949) p.1232

[7] Van Neumann, J., and Goldstine, H.H., Bull. Amer. Math. Soc. vol. 53, (1947), p.1021

[8] Kerns, D.M., Analysis of Symmetrical Waveguide Junctions, U.S. Nat'l. Bur. of Stds. Jl. Res. vol. 46 (1951) p.267-282